

Solving SDP Completely with an Interior Point Oracle

Bruno F. Lourenço^{*} Masakazu Muramatsu[†] Takashi Tsuchiya[‡]

July 2015 (Revised: November 2015)

Abstract

We suppose the existence of an oracle which is able to solve any semidefinite programming (SDP) problem having interior feasible points at both its primal and dual sides. We note that such an oracle might not be able to directly solve general SDPs even after certain regularization schemes such as facial reduction or Ramana's dual are applied. The main objective of this work is to fill this gap and to show in detail how one can use such an oracle to '*completely solve*' an arbitrary SDP. Here, we use the term *to completely solve* an SDP to mean a scheme which works in the following way; given an SDP, the scheme checks whether it is feasible or not, and whenever feasible, computes its optimal value, and if the optimal value is attained, obtains a maximal rank optimal solution. If the optimal value is not attained, computes a feasible solution whose objective value is arbitrarily close to the optimal value. Moreover, if the original SDP is infeasible, it distinguishes between strong and weak infeasibility, and in the case of weak infeasibility, can compute a point in the corresponding affine space whose distance to the positive semidefinite cone is less than an arbitrary small given positive number.

1 Introduction

Consider the following pair of primal and dual conic linear programs.

$$\begin{aligned} \inf_x \quad & \langle c, x \rangle \\ \text{subject to} \quad & \mathcal{A}x = b \\ & x \in \mathcal{K} \end{aligned} \tag{P}$$

$$\begin{aligned} \sup_y \quad & \langle b, y \rangle \\ \text{subject to} \quad & c - \mathcal{A}^T y \in \mathcal{K}^*, \end{aligned} \tag{D}$$

where $\mathcal{K} \subseteq \mathbb{R}^n$ is a closed convex cone and \mathcal{K}^* is the dual cone $\{s \in \mathbb{R}^n \mid \langle s, x \rangle \geq 0, \forall x \in \mathcal{K}\}$. We have that $\mathcal{A} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a linear map, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ and \mathcal{A}^T denotes the adjoint map. We also have $\mathcal{A}^T y = \sum_{i=1}^m \mathcal{A}_i y_i$, for certain elements $\mathcal{A}_i \in \mathbb{R}^n$. The inner product is denoted by $\langle \cdot, \cdot \rangle$. We will also use the notation $(\mathcal{K}, \mathcal{A}, b, c)$ to refer to the pair (P), (D).

In this article, we are mainly interested in the case where $\mathcal{K} = \mathcal{S}_+^n$, where \mathcal{S}_+^n is the cone of $n \times n$ symmetric positive semidefinite matrices. In this case, \mathcal{A} is a map from \mathcal{S}^n to \mathbb{R}^m , where \mathcal{S}^n is the space of $n \times n$ symmetric matrices and $\langle c, x \rangle = \text{trace}(cx)$.

^{*}Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, 2-12-1-W8-41 Ookayama, Meguro-ku, Tokyo 152-8552, Japan. (E-mail: flourenco.b.aa@m.titech.ac.jp)

[†]Department of Computer Science, The University of Electro-Communications 1-5-1 Chofugaoka, Chofu-shi, Tokyo, 182-8585 Japan. (E-mail: muramatsu@cs.uec.ac.jp)

[‡]National Graduate Institute for Policy Studies 7-22-1 Roppongi, Minato-ku, Tokyo 106-8677, Japan. (E-mail: tsuchiya@grips.ac.jp)

M. Muramatsu and T. Tsuchiya are supported in part with Grant-in-Aid for Scientific Research (B)24310112 and (C) 26330025. M. Muramatsu is also partially supported by the Grant-in-Aid for Scientific Research (B)26280005. T. Tsuchiya is also partially supported by the Grant-in-Aid for Scientific Research (B)15H02968.

Interior point methods (IPMs) [15, 2] are one of the standard methods for solving semidefinite programs $(\mathcal{S}_+^n, \mathcal{A}, b, c)$. However, in order to function properly, they require that the problem at hand satisfy certain regularity conditions which may fail to be satisfied and this leads to numerical difficulties. The usual requirement is that there is a primal feasible solution x such that x is positive definite *and* that there is a dual feasible solution y such that the corresponding slack $s = c - \mathcal{A}^T y$ is positive definite. In other words, Slater's condition must hold for both (P) and (D). In this article, we suppose the existence of an idealized machine that is able to solve any SDP having primal and dual interior feasible points and discuss whether it could be used to solve other SDPs which might not satisfy that assumption. In this paper, we use the expression that an algorithm or a scheme *completely solves* SDP when it works in the following way.

1. It checks whether the SDP is feasible or not.
2. When the SDP is feasible, it computes the optimal value. If the optimal value is attained, it computes a maximal rank optimal solution. Moreover, if the optimal value is not attained, given arbitrary small $\epsilon > 0$, it can compute a feasible point whose objective value is within ϵ distance of the optimal value.
3. When the SDP is infeasible, it distinguishes whether it is strongly infeasible, or weakly infeasible. Whenever the SDP is strongly infeasible, it computes a certificate of infeasibility. If the SDP is weakly infeasible, then it can compute a point in the corresponding affine space whose distance to the positive semidefinite cone is less than an arbitrary small given positive number.

First, we will explain why this is a non-trivial task while connecting to previous research on conic linear programming.

- *Interior point algorithms and software.* Most modern IPM softwares [26, 8, 28] do not require explicit knowledge of an interior feasible point beforehand. SeDuMi [26], for instance, transforms a standard form problem into the so-called homogeneous self-dual formulation, which has a trivial starting point. SDPA [8] and SDPT3 [28] use an infeasible interior point method. The fact that these methods can work without explicit knowledge of an interior feasible point, does not mean that they do *not require the existence of an interior feasible point*. Quite the opposite, the absence of interior feasible points may introduce theoretical and numerical difficulties in recovering a solution for the original problem. Also, detection of infeasibility is a complicated task. Some interior point methods, such as the one discussed in [16] by Nesterov, Todd and Ye, are able to obtain a certificate of infeasibility if the problem is dual or primal strongly infeasible, but the situation is less clear in the presence of weak infeasibility.
- *Ramana's extended dual.* When neither the primal nor the dual have interior feasible points, there could be a nonzero duality gap between the (P) and (D). To address this, Ramana [23, 24] developed an alternative dual for (D) that requires no regularity assumptions for the case $\mathcal{K} = \mathcal{S}_+^n$. Remarkable features of Ramana's dual include the fact that it can be written as a SDP, always affords zero duality gap and that the dual is always attained whenever the primal optimal value is finite. However, Ramana's dual is not necessarily suitable to be used with IPMs due to the fact that it does not ensure the existence of interior feasible points at both sides.
- *Facial reduction.* Let \mathcal{F}_{\min}^D be the minimal face of \mathcal{S}_+^n which contains the feasible slacks of (D). If we substitute \mathcal{S}_+^n for \mathcal{F}_{\min}^D in (D), then (D) will have a relative interior slack. This means that if the (Lagrangian) primal problem is computed with respect to \mathcal{F}_{\min}^D instead of \mathcal{S}_+^n , Slater's condition will be satisfied and there will be no duality gap. The process of finding \mathcal{F}_{\min}^D is called *facial reduction* [31, 19] and was developed originally by Borwein and Wolkowicz [5, 4]. However, one important point is that although there is a dual relative interior, there is no guarantee that there will be a primal relative interior feasible point. So, again, even this regularized problem might fail to have interior solutions at both sides.

There is a growing body of research aimed at understanding SDPs having pathological behaviours such as nonzero duality gaps and weak infeasibility. Here we will mention a few of them. A problem is said to be weakly infeasible if there is no feasible solution but the distance between the underlying affine space and the

cone under consideration is zero. Weak infeasibility is known to be very hard to detect numerically, see for instance Pólik and Terlaky [21]. In [30], Waki showed that weakly infeasible problems sometimes arise from polynomial optimization. There is also a discussion on weak infeasibility in the context of SDPs in [12], where the authors introduce the notion of hyper feasible partitions and show that for weakly infeasible problems at most $n - 1$ directions are needed to construct points close to \mathcal{S}_+^n and that belong to the underlying affine space defined by the problem. Hyper feasible partitions have many interesting properties and they can be used to find smaller problems that almost preserve the feasibility status of a given problem, see Theorem 12. The result on the number of directions was recently generalized to arbitrary closed convex cones by Liu and Pataki, see [10] for more details.

It is hard to obtain finite certificates of infeasibility for SDPs, because Farkas' Lemma does not hold in general for non-polyhedral cones. The first finite certificate was obtained by Ramana in [23]. Since then, Sturm mentioned the possibility of obtaining a finite certificate for infeasibility by using the directions produced in his regularization procedure, see page 1243 of [27]. More recently, Liu and Pataki have also obtained finite certificates through elementary reformulations [11]. Interestingly, Klep and Schweighofer [9] also obtained certificates through a completely different approach using tools from real algebraic geometry.

In [32], Waki, Nakata and Muramatsu discussed SDP instances for which known solvers failed to obtain the correct answer and in one case, this happened even though the problem had an interior feasible point at the primal side. In [18], Pataki gave a definition of “bad behaviour” and showed that all SDPs in that class can be put in the same form, after performing an elementary reformulation. A discussion on duality gaps and many interesting examples of pathological SDPs are given by Tunçel and Wolkowicz in [29].

The fact is that, in general, SDPs can display several nasty behaviors and this stands in sharp contrast to the relatively nice nature of SDPs having primal and dual interior feasible points. The main goal of this paper is to show that a SDP can be completely solved by considering a finite sequence of very well behaved problems. In particular, we demonstrate that it is possible to further regularize a SDP after facial reduction is done. By doing so, we can convert a strongly feasible SDP with possibly weakly feasible dual to another one satisfying strong feasibility at both sides without changing the optimal value. This is one of the main contribution of the paper and is treated in Section 5. We also make it clear that facial reduction can be performed by solving nice SDPs with primal and dual interior feasible points. This was already discussed by Cheung, Schurr and Wolkowicz [6], but it is a point which had not been emphasized much in the literature. In our context, this facet of facial reduction is critically important. Therefore, for the sake of self-containment, we deal with this issue in more detail in Section 3.

This work is organized as follows. In Section 2, we review a few basic notions and also facial reduction. In Section 3 we define the interior point oracle and discuss a version of facial reduction for SDPs. In Section 4, we recall some properties of hyper feasible partitions and show how they can be constructed with the interior point oracle. Section 5 shows how to compute the optimal value and in Section 6, we discuss how to check if the value is attained or not. Section 7 shows a complete list of steps for completely solving (D) and in Section 8 we compare our approach with previous works. Section 9 wraps up this paper.

2 Preliminary discussion and review of relevant notions

Let C be a closed convex set. Its relative interior, closure and relative boundary are denoted by $\text{ri}C$, $\text{cl}C$ and $\text{relbd}C$, respectively. For \mathcal{K} a closed convex cone, we denote by $\text{lin}\mathcal{K}$ the largest subspace contained in \mathcal{K} . A convex subset \mathcal{F} of \mathcal{K} is said to be a face of \mathcal{K} if the condition “ $a, b \in \mathcal{K}, \frac{a+b}{2} \in \mathcal{F}$ ” implies $a, b \in \mathcal{F}$. Given a conic linear program $(\mathcal{K}, \mathcal{A}, b, c)$, we will denote by θ_P and θ_D , the primal and dual optimal values, respectively. It is understood that $\theta_P = +\infty$ if (P) is infeasible and $\theta_D = -\infty$ if (D) is infeasible. The pair (P) and (D) is said to have *zero duality gap* if $\theta_P = \theta_D$. The primal and dual feasible regions are defined as follows:

$$\begin{aligned}\mathcal{F}_P &= \{x \in \mathbb{R}^n \mid \mathcal{A}x = b, x \in \mathcal{K}\} \\ \mathcal{F}_D &= \{y \in \mathbb{R}^m \mid c - \mathcal{A}^T y \in \mathcal{K}^*\} \\ \mathcal{F}_D^S &= \{s \in \mathcal{K}^* \mid \exists y \in \mathbb{R}^m, s = c - \mathcal{A}^T y\},\end{aligned}$$

where an element $s \in \mathcal{F}_D^S$ is said to be a *dual slack*. The dual optimal value θ_D is said to be attained if there is a dual feasible $y \in \mathcal{F}_D$ such that $\langle b, y \rangle = \theta_D$. The notion of primal attainment is analogous. We recall the following basic constraint qualification.

Proposition 1 (Slater). *Let $(\mathcal{K}, \mathcal{A}, b, c)$ be a conic linear program.*

- i. If there exists $x \in \text{ri } \mathcal{K} \cap \mathcal{F}_P$, then $\theta_P = \theta_D$. In addition, if $\theta_P > -\infty$ holds as well then the dual optimal value is attained.*
- ii. If there exists $s \in \mathcal{F}_D^S \cap \text{ri } \mathcal{K}^*$, then $\theta_P = \theta_D$. In addition, if $\theta_D < +\infty$ holds as well, the primal optimal value is attained.*

For the reader's convenience, before we proceed we recall a few basic facts from convex analysis.

Lemma 2. *Let \mathcal{K} be a closed convex cone, $\mathcal{I} \in \text{ri } \mathcal{K}$, $x \in \mathcal{K}$ and $z \in \mathcal{K}^*$.*

- i. $\mathcal{K}^{**} = \mathcal{K}$.*
- ii. $\text{lin } \mathcal{K} = \mathcal{K}^{\perp}$, where \mathcal{K}^{\perp} is a short-hand for $(\mathcal{K}^*)^{\perp}$.*
- iii. $x + \mathcal{I} \in \text{ri } \mathcal{K}$.*
- iv. There exists $\alpha > 1$ such that $\alpha \mathcal{I} + (1 - \alpha)x \in \mathcal{K}$.*
- v. $z \in \mathcal{K}^{\perp}$ if and only if $\langle \mathcal{I}, z \rangle = 0$.*

Proof. *i.* This is the famous bipolar theorem, see Theorem 14.1 of [25].

ii. If $z \in \text{lin } \mathcal{K}$, then $\langle z, y \rangle \geq 0$ and $\langle -z, y \rangle \geq 0$, for every $y \in \mathcal{K}^*$. It follows that $z \in \mathcal{K}^{\perp}$. Reciprocally, if $z \in \mathcal{K}^{\perp}$, then $z \in \mathcal{K}^{**} = \mathcal{K}$, by the bipolar theorem. Since \mathcal{K}^{\perp} is a subspace, we have $\mathcal{K}^{\perp} \subseteq \text{lin } \mathcal{K}$.

iii. Since $\mathcal{I} \in \text{ri } \mathcal{K}$, for any $z \in \mathcal{K}$ we have that all points in the relative interior of the line segment connecting z and \mathcal{I} also belong to the relative interior of \mathcal{K} , see Theorem 6.1 of [25]. Since $x + \mathcal{I} = \mathcal{I} \frac{1}{2} + (2x + \mathcal{I}) \frac{1}{2}$, we have $x + \mathcal{I} \in \text{ri } \mathcal{K}$.

iv. See Theorem 6.4 of [25].

v. If $z \in \mathcal{K}^{\perp}$, it is clear that $\langle \mathcal{I}, z \rangle$ is zero. Now, suppose that $\langle \mathcal{I}, z \rangle$ is zero. By item *iv*, there is $\alpha > 1$ such that $\alpha \mathcal{I} + (1 - \alpha)x \in \mathcal{K}$. On one hand, since $z \in \mathcal{K}^*$, we have $\langle u, z \rangle \geq 0$. On the other, $\langle u, z \rangle = (1 - \alpha)\langle x, z \rangle \leq 0$. So, we must have $\langle x, z \rangle = 0$. As x is an arbitrary element, it holds that $z \in \mathcal{K}^{\perp}$. □

It is a consequence of classical separation theorems that existence of a primal relative interior feasible point can be translated to a statement about the dual problem. We recall that if C_1 and C_2 are two non-empty convex sets then $\text{ri } C_1 \cap \text{ri } C_2 = \emptyset$ if and only if C_1 and C_2 can be *properly separated* [25, Theorem 11.3]. Two convex sets are said to be properly separated if there is some hyperplane such that the sets are contained in opposite closed half-spaces and at least one of them is not entirely contained in the hyperplane. Moreover, if C_1 is polyhedral and $C_1 \cap \text{ri } C_2 = \emptyset$ we can ensure that the separation can be done in such a way that C_2 is not contained in the hyperplane [25, Theorem 20.2]. This leads to the following fact, which can be used as an indirect way of showing existence of relative interior feasible points, see also Lemma 3.2 in [31].

Proposition 3. *Let $(\mathcal{K}, \mathcal{A}, b, c)$ be a conic linear program.*

- i. Suppose that there exists some \hat{x} (not necessarily feasible) such that $\mathcal{A}\hat{x} = b$. Then, $\text{ri } \mathcal{K} \cap \mathcal{F}_P \neq \emptyset$ if and only if the following conditions hold at the same time:*

$$\langle b, y \rangle = 0 \quad \text{and} \quad -\mathcal{A}^T y \in \mathcal{K}^* \Rightarrow -\mathcal{A}^T y \in \mathcal{K}^{\perp}. \quad (1)$$

$$\{y \mid \langle b, y \rangle > 0, -\mathcal{A}^T y \in \mathcal{K}^*\} = \emptyset \quad (2)$$

ii. The condition $\text{ri } \mathcal{K}^* \cap \mathcal{F}_D^S \neq \emptyset$ holds if and only if the following two conditions hold at the same time:

$$\langle c, x \rangle = 0 \quad \text{and} \quad \mathcal{A}x = 0 \quad \text{and} \quad x \in \mathcal{K} \Rightarrow x \in \text{lin } \mathcal{K}. \quad (3)$$

$$\{x \mid \langle c, x \rangle < 0, \mathcal{A}x = 0, x \in \mathcal{K}\} = \emptyset \quad (4)$$

Proof. i. We first show that if Conditions (1) or (2) fail to hold then (P) cannot have a relative interior solution. If the latter fails, then (P) is infeasible. Now suppose that the former fails to hold and let y be such that $\langle b, y \rangle = 0$, $-\mathcal{A}^T y \in \mathcal{K}^*$ but $-\mathcal{A}^T y \notin \mathcal{K}^\perp$. Then, if $x \in \mathcal{F}_P$, we have $\langle -\mathcal{A}^T y, x \rangle = -\langle y, b \rangle = 0$. Since $-\mathcal{A}^T y \notin \mathcal{K}^\perp$, x cannot be a relative interior point, due to item v. of Lemma 2.

To conclude, we show that if $\text{ri } \mathcal{K} \cap \mathcal{F}_P = \emptyset$, then at least one of (1) or (2) must fail. The condition $\text{ri } \mathcal{K} \cap \mathcal{F}_P = \emptyset$ holds if and only if the nonempty polyhedral set $C = \{x \mid \mathcal{A}x = b\}$ and \mathcal{K} can be properly separated. Proper separation implies the existence of a nonzero $s \in \mathbb{R}^n$ and $\gamma \in \mathbb{R}$ such that

$$\langle s, \hat{x} + d \rangle \leq \gamma \leq \langle s, k \rangle, \quad (5)$$

for every $d \in \ker \mathcal{A}$ and $k \in \mathcal{K}$. The only way (5) can hold is if $s \in \mathcal{K}^* \cap (\ker \mathcal{A})^\perp$ and $\gamma \leq 0$. Moreover, the fact that C is polyhedral and Theorem 20.2 of [25] imply that we can also assume that \mathcal{K} is not contained in the hyperplane $\{x \mid \langle s, x \rangle = \gamma\}$. Since $s \in (\ker \mathcal{A})^\perp$, there exists some $y \in \mathbb{R}^m$ such that $s = -\mathcal{A}^T y$. We then have $\langle s, \hat{x} \rangle = -\langle b, y \rangle \leq 0$. We then have two cases. If $-\langle b, y \rangle = 0$, then $\gamma = 0$, which implies that $s \notin \mathcal{K}^\perp$, in order for the separation to be proper. This implies that Condition (1) fails to hold. On the other hand, if $-\langle b, y \rangle < 0$, then Condition (2) fails.

ii. The dual proof is analogous. The condition $\text{ri } \mathcal{K}^* \cap \mathcal{F}_D^S = \emptyset$ holds if and only if the polyhedral set $C = \{c - \mathcal{A}^T y \mid y \in \mathbb{R}^m\}$ and \mathcal{K}^* can be properly separated. This implies the existence of x such that $\mathcal{A}x = 0$, $x \in \mathcal{K}^{**}$ and either: i. $\langle c, x \rangle = 0$ and $x \notin \mathcal{K}^{*\perp}$; or ii. $\langle c, x \rangle < 0$. We also recall that $\mathcal{K}^{**} = \mathcal{K}$ and $\mathcal{K}^{*\perp} = \text{lin } \mathcal{K}$, by items i. and ii. of Lemma 2. Thus we can conclude that (3) and (4) are equivalent to the existence of a dual relative interior slack. \square

Sometimes, we will be interested solely in the conic feasibility problem, which we will denote by $(\mathcal{K}, \mathcal{L}, c)$. This is the problem of seeking a point in the intersection $(\mathcal{L} + c) \cap \mathcal{K}$, where $\mathcal{L} \subseteq \mathbb{R}^n$ is a vector subspace and $c \in \mathbb{R}^n$. There are only four mutually exclusive categories that $(\mathcal{K}, \mathcal{L}, c)$ can fall in:

- i. strong feasibility: $\text{ri } \mathcal{K} \cap (\mathcal{L} + c) \neq \emptyset$,
- ii. weak feasibility: $\text{ri } \mathcal{K} \cap (\mathcal{L} + c) = \emptyset$, but $\mathcal{K} \cap (\mathcal{L} + c) \neq \emptyset$.
- iii. weak infeasibility: $\mathcal{K} \cap (\mathcal{L} + c) = \emptyset$, but the Euclidean distance $\text{dist}(\mathcal{L} + c, \mathcal{K}) = 0$,
- iv. strong infeasibility: $\text{dist}(\mathcal{L} + c, \mathcal{K}) > 0$.

In view of the above categories, the usual assumption underlying interior point methods amounts to requiring both primal and dual strong feasibility. This is also equivalent to, for instance, the existence of a primal interior feasible point together with boundness of the level sets of the primal objective function, see Proposition 3. As a convention, if a feasibility problem is formulated as the problem of finding a x satisfying both $\mathcal{A}x = b$ and $x \in \mathcal{K}$ but the linear system $\mathcal{A}x = b$ has no solution, we will say that it is strongly infeasible. We have the following well-known characterization of strong infeasibility.

Proposition 4. i. The primal (P) is strongly infeasible if and only if there is no solution to the linear system $\mathcal{A}x = b$ or if there exists y such that

$$\langle b, y \rangle = 1 \quad \text{and} \quad -\mathcal{A}^T y \in \mathcal{K}^*. \quad (6)$$

ii. The dual (D) is strongly infeasible if and only if there exists x such that

$$\langle c, x \rangle = -1 \quad \text{and} \quad \mathcal{A}x = 0 \quad \text{and} \quad x \in \mathcal{K} \quad (7)$$

Proof. It is a consequence of Theorem 11.4 in [25], which states that two convex sets C_1, C_2 can be strongly separated if and only if the Euclidean distance between them is positive. Strong separation means that there exists a separating hyperplane H and $\epsilon > 0$ such that $C_1 + \epsilon B$ and $C_2 + \epsilon B$ lie in opposite open half-spaces, where B is the unit ball. One then proceeds as in the proof of 3. See also Lemma 5 in [13]. \square

We also remark the following immediate consequence of Proposition 4.

Corollary 5. *i. If the primal (P) is strongly infeasible and the dual (D) is feasible, then $\theta_D = +\infty$.*

ii. If the dual is strongly infeasible and the primal is feasible then $\theta_P = -\infty$.

Proof. To prove item *i.*, note that by Proposition 4, strong infeasibility at the primal side implies the existence of y such that $\langle b, y \rangle = 1$ and $-\mathcal{A}^T y \in \mathcal{K}^*$. So, if \hat{y} is any dual feasible solution then $\hat{y} + \alpha y$ is also feasible for every $\alpha \geq 0$ and we can make the objective function value as large as we want. The proof of item *ii.* is analogous. \square

2.1 Facial Reduction

In this section, we will briefly discuss facial reduction. Here, we will focus on problems formulated in the dual form (D). But it is clear that any analysis carried out for (D) can be translated back to (P).

Whenever (D) lacks a relative interior solution, we may reformulate (D) over a lower dimensional face of \mathcal{K}^* . The basis for this idea is contained in Proposition 3: the absence of a dual interior solution is equivalent to the existence of an x satisfying $\langle c, x \rangle \leq 0$ and $x \in \ker \mathcal{A} \cap \mathcal{K}$ such that one of the two alternatives is satisfied: *i.* $\langle c, x \rangle < 0$, in which case (D) is infeasible; *ii.* $\langle c, x \rangle = 0$ and $x \notin \mathcal{K}^{*\perp} = \text{lin } \mathcal{K}$, in which case $\mathcal{F}_D^S \subseteq \mathcal{K}^* \cap \{x\}^\perp \subsetneq \mathcal{K}^*$, because $\langle c, x \rangle = \langle c - \mathcal{A}^T y, x \rangle = 0$ for $y \in \mathcal{F}_D$. Note that $\mathcal{F}_2 = \mathcal{K}^* \cap \{x\}^\perp$ is a proper face of \mathcal{K}^* , since $x \notin \text{lin } \mathcal{K}$. We can then reformulate (D) as a problem over \mathcal{F}_2 , that is, we consider the problem $\sup\{\langle b, y \rangle \mid c - \mathcal{A}^T y \in \mathcal{F}_2\}$. It is clear that as long as $\text{ri } \mathcal{F}_2 \cap (\mathcal{F}_D^S) = \emptyset$, we can repeat this process and either descend to a smaller face of \mathcal{K}^* or declare infeasibility. After a few iterations, we will end with either some face \mathcal{F}_ℓ such that $\text{ri } \mathcal{F}_\ell \cap (\mathcal{F}_D^S) \neq \emptyset$ or we will eventually find out that the problem is infeasible, in this case $\mathcal{F}_\ell = \emptyset$. Note that \mathcal{F}_ℓ must be the smallest face \mathcal{F}_{\min}^D of \mathcal{K}^* which contains \mathcal{F}_D^S . This process is called *facial reduction* and it aims at finding \mathcal{F}_{\min}^D . The direction x will be henceforth called a *reducing direction*. If there is no dual feasible solution, we have $\mathcal{F}_{\min}^D = \emptyset$. The minimal face also has the following well-known characterization, see for instance, Proposition 3.2.2 in [17].

Proposition 6. *Let \mathcal{F} be a face of \mathcal{K}^* containing \mathcal{F}_D^S . Then the conditions below are equivalent.*

- i.* $\mathcal{F}_D^S \cap \text{ri } \mathcal{F} \neq \emptyset$.
- ii.* $\text{ri } \mathcal{F}_D^S \subseteq \text{ri } \mathcal{F}$.
- iii.* $\mathcal{F} = \mathcal{F}_{\min}^D$.

Facial reduction is a very powerful procedure and it can be used to solve feasibility problems over arbitrary closed convex cones. What is troublesome about it is that searching for x could be, in principle, as hard as solving the original problem itself. However, an important point is that it is possible to carefully formulate the problem of seeking x as a linear conic program having relative interior feasible points at both its primal and dual sides, as we will do in Section 3. Another alternative is to relax the search criteria in order to make the problem of finding x more tractable at the cost of, perhaps, settling for a face other than \mathcal{F}_{\min}^D , as in the Partial Facial Reduction approach of Permenter and Parrilo [20].

One of the problems with facial reduction is that even if we find some face for which the reduced system is strongly feasible, there is no guarantee that the corresponding primal will also be strongly feasible. That is, facial reduction only restores strong feasibility at one of the sides of the problem. So the unfortunate fact remains that even after performing facial reduction one may end up with a problem that is still not suitable to be solved by primal-dual interior point methods.

We remark that strong feasibility at only one of the sides of the problem can also be a source of numerical trouble. In Section 2 of [32], Waki, Nakata and Muramatsu show an example of a SDP that is primal strongly

feasible but dual weakly feasible. Its optimal value is zero but both SDPA [8] and SeDuMi [26] output 1 instead.

2.2 Facial structure of \mathcal{S}_+^n

The cone of positive semidefinite symmetric matrices has a very special structure and every face of \mathcal{S}_+^n is linearly isomorphic to some \mathcal{S}_+^r for $r \leq n$. The following proposition is well-known, but we give a short proof here.

Proposition 7. *Let \mathcal{F} be a nonempty face of \mathcal{S}_+^n . Then:*

- i. *For all $x \in \text{ri } \mathcal{F}$ and $y \in \mathcal{F}$, we have $\ker x \subseteq \ker y$.*
- ii. *There exists $r \leq n$ and an orthogonal $n \times n$ matrix q such that*

$$q^T \mathcal{F} q = \left\{ \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \mid a \in \mathcal{S}_+^r \right\} \quad (8)$$

Proof. i. Since x is a relative interior feasible point, by item iv of Lemma 2, there is $\alpha > 1$ such that $z = \alpha x + (1 - \alpha)y \in \mathcal{F}$. Let $d \in \ker x$. Then, $\langle zd, d \rangle = (1 - \alpha)\langle yd, d \rangle \leq 0$. However, z is positive semidefinite, so we must have $\langle yd, d \rangle = 0$ which implies $d \in \ker y$.

- ii. Let $x \in \text{ri } \mathcal{F}$ and $r = n - \dim \ker x$. Let q be an orthogonal matrix such that the last $\dim \ker x$ columns span the kernel of x . Using the previous item and the fact that the matrices in \mathcal{F} are positive semidefinite, it is possible to conclude that Equation (8) holds. □

It follows that if \mathcal{F} is a nonempty face, the rank of the elements in its relative interior is constant and they share the same kernel. We will call this quantity the *rank* of \mathcal{F} . Using the fact that q is orthogonal, we have that the dual of \mathcal{F} satisfies

$$q^T \mathcal{F}^* q = (q^T \mathcal{F} q)^* = \left\{ \begin{pmatrix} a & b \\ b & d \end{pmatrix} \in \mathcal{S}^n \mid a \in \mathcal{S}_+^r \right\}. \quad (9)$$

3 Definition of oracle and related discussion

We assume the existence of the following oracle.

Definition 8 (The Interior Point Oracle). *Let $\mathcal{A} : \mathcal{S}^n \rightarrow \mathbb{R}^m$ be a linear map, $b \in \mathbb{R}^m$ and $c \in \mathcal{S}^n$ such that $(\mathcal{S}_+^n, \mathcal{A}, b, c)$ is both primal strongly feasible and dual strongly feasible. The interior point oracle receives as input $(\mathcal{S}_+^n, \mathcal{A}, b, c)$ and outputs a pair (x^*, y^*) such that x^* is a primal optimal solution and y^* is a dual optimal solution.*

We will refer to the oracle in Definition 8 as (IPO). We can regard (IPO) as a machine running an idealized version of either the homogeneous self-dual embedding [22, 7, 14], an infeasible interior point method [16] or even the ellipsoid method. The machine does not need to receive interior points, but it is only guaranteed to work properly if both (P) and (D) have interior feasible points. We emphasize that this assumption is essentially equivalent to requiring existence of an interior feasible point at one side of the problem, together with boundedness of level set of the objective function at that same side, by Proposition 3. Note that no assumption is made on the inner workings of the oracle.

Our discussion could be framed precisely in the real computation model of Blum, Shub and Smale [3], but for simplicity, apart from (IPO), we will only assume that all the elementary real arithmetic operations can be carried out exactly. Under this setting, we can use (IPO) to compute the square root of any positive real number β by solving the SDP $\theta_D = \sup\{x \mid \begin{pmatrix} 1 & x \\ x & \beta \end{pmatrix} \in \mathcal{S}_+^2\}$. Since it has both primal and dual interior feasible solutions, it is licit to call (IPO) to solve it. Another technical point we should mention is that in

the definition of the oracle, the affine space is contained in the space of $n \times n$ symmetric matrices and the optimization is carried over \mathcal{S}_+^n . Note that n is the same for both \mathcal{S}_+^n and \mathcal{S}^n . However, for fixed \mathcal{A}, b, c we might be interested in solving problems over a face \mathcal{F} of \mathcal{S}^n or over its dual \mathcal{F}^* . Even if $(\mathcal{F}, \mathcal{A}, b, c)$ is both primal and dual strongly feasible, it is not immediately clear how to use (IPO) to solve it, since $(\mathcal{F}, \mathcal{A}, b, c)$ is not exactly a SDP. One possibility would be to consider a, *a priori*, stronger oracle that is also able to solve strongly feasible problems over faces of \mathcal{S}_+^n . The next proposition shows that there is no need for that.

Proposition 9. *Let $\mathcal{A} : \mathcal{S}^n \rightarrow \mathbb{R}^m$ be a linear map, $b \in \mathbb{R}^m$, $c \in \mathcal{S}^n$ and \mathcal{F} be a non-empty proper face of \mathcal{S}_+^n . Suppose that $(\mathcal{F}, \mathcal{A}, b, c)$ is both primal and dual strongly feasible. That is, there is $x \in \mathcal{S}^n$ satisfying $\mathcal{A}x = b$ and $x \in \text{ri } \mathcal{F}$ and $y \in \mathbb{R}^m$ such that $c - \mathcal{A}^T y \in \text{ri } \mathcal{F}^*$.*

Then, it is possible to transform $(\mathcal{F}, \mathcal{A}, b, c)$ into an equivalent auxiliary problem that is solvable by (IPO). Moreover, a pair of primal and dual optimal solutions is readily recoverable from the optimal solutions of the auxiliary problem.

Proof. We defer the proof until the Appendix A. □

For convenience, we will analyze SDPs in dual format, but it is always possible to translate to primal format, if necessary. We will accomplish the following with (IPO).

- i. Find the minimal face \mathcal{F}_{\min}^D of \mathcal{S}_+^n which contains the feasible region of (D). If $\mathcal{F}_{\min}^D = \emptyset$, distinguish between weak and strong infeasibility. For weakly infeasible problems, (IPO) can be used for finding (infeasible) slacks arbitrarily close to \mathcal{S}_+^n .
- ii. Decide whether the optimal value θ_D is attained or not.
- iii. In case of attainment, find an optimal solution of maximal rank. If the optimal value is unattained, then (IPO) can be used to find points in \mathcal{F}_D^S that are arbitrarily close to optimality.

In this section, we will focus on finding \mathcal{F}_{\min}^D and we will resort to Facial Reduction to accomplish this. The next lemma shows how to formulate the problem of finding a reducing direction as a problem that can be fed to (IPO). We remark that in [6], Cheung, Schurr and Wolkowicz also discuss an auxiliary problem that is primal and dual strongly feasible, see the problem (AP) therein. However, (AP) uses an additional second order cone constraint. Since a second order cone constraint can be transformed to a semidefinite constraint, (AP) is also an SDP. However, (AP) is not readily generalizable to other families of cones that are not able to express those constraints.

Lemma 10. *Let \mathcal{F} be a non-empty face of \mathcal{S}_+^n containing \mathcal{F}_D^S and let $\mathcal{I} \in \text{ri } \mathcal{F}$, $\mathcal{I}^* \in \text{ri } \mathcal{F}^*$. Now consider the following pair of primal and dual problems.*

$$\begin{array}{ll} \underset{x, t, w}{\text{minimize}} & t \end{array} \tag{P_{\mathcal{F}}}$$

$$\text{subject to} \quad -\langle c, x - t\mathcal{I}^* \rangle + t - w = 0 \tag{10}$$

$$\langle \mathcal{I}, x \rangle + w = 1 \tag{11}$$

$$\mathcal{A}x - t\mathcal{A}\mathcal{I}^* = 0$$

$$(x, t, w) \in \mathcal{F}^* \times \mathcal{S}_+^1 \times \mathcal{S}_+^1$$

$$\begin{array}{ll} \underset{y_1, y_2, y_3}{\text{maximize}} & y_2 \end{array} \tag{D_{\mathcal{F}}}$$

$$\text{subject to} \quad cy_1 - \mathcal{I}y_2 - \mathcal{A}^T y_3 \in \mathcal{F} \tag{12}$$

$$1 - y_1(1 + \langle c, \mathcal{I}^* \rangle) + \langle \mathcal{I}^*, \mathcal{A}^T y_3 \rangle \geq 0$$

$$y_1 - y_2 \geq 0 \tag{13}$$

The following properties hold.

- i. Both $(P_{\mathcal{F}})$ and $(D_{\mathcal{F}})$ have relative interior feasible points (so we may invoke (IPO) to solve it, by Proposition 9).
- ii. Let (x^*, t^*, w^*) be a primal optimal solution. The optimal value is zero if and only if $\mathcal{F}_{\min}^D \subsetneq \mathcal{F}$. Moreover, if the optimal value is zero, we have $\langle c, x^* \rangle < 0$ and $\mathcal{F}_D^S = \mathcal{F}_{\min}^D = \emptyset$, or $\langle c, x^* \rangle = 0$ and $\mathcal{F}_D^S \subseteq \mathcal{F} \cap \{x^*\}^\perp \subsetneq \mathcal{F}$.
- iii. Let (y_1^*, y_2^*, y_3^*) be a dual optimal solution. If the common optimal value is nonzero, then $\mathcal{F}_{\min}^D = \mathcal{F}$ and $s = c - \mathcal{A}^T \frac{y_3^*}{y_1^*}$ is a dual optimal solution satisfying $s \in (\text{ri } \mathcal{F}_D^S) \cap \text{ri } \mathcal{F}$.

Note that Proposition 7 together Equations (8) and (9) indicate how to find \mathcal{I} and \mathcal{I}^* without much effort. One can take $\mathcal{I} = \mathcal{I}^*$ and let \mathcal{I} be such that $q^T \mathcal{I} q = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}$, where a is the $r \times r$ identity matrix.

Proof. i. Let $t = \frac{1}{\langle \mathcal{I}, \mathcal{I}^* \rangle + 1}$, $w = \frac{1}{\langle \mathcal{I}, \mathcal{I}^* \rangle + 1}$ and $x = \frac{\mathcal{I}^*}{\langle \mathcal{I}, \mathcal{I}^* \rangle + 1}$. Then (x, t, w) is an interior solution to $(P_{\mathcal{F}})$. To show that $(D_{\mathcal{F}})$ has a relative interior solution, we use just observe that $(0, -1, 0)$ is a dual feasible solution which corresponds to relative interior slack.

- ii. Now, let (x^*, t^*, w^*) be an optimal solution. If $t^* = 0$, then $x^* \in \ker \mathcal{A} \cap \mathcal{F}^*$ and Equation (10) implies that $\langle c, x^* \rangle \leq 0$. If y is a dual feasible solution for (D), we have $c - \mathcal{A}^T y \in \mathcal{F}$, so that $\langle c - \mathcal{A}^T y, x^* \rangle = \langle c, x^* \rangle \leq 0$. If $\langle c, x^* \rangle < 0$, then it must be the case that $\mathcal{F}_D^S = \emptyset$. If $\langle c, x^* \rangle = 0$, then $c - \mathcal{A}^T y \in \{x^*\}^\perp$. Moreover, Equation (10) implies that $w^* = 0$ as well. Using Equation (11), we obtain $\langle \mathcal{I}, x^* \rangle = 1$. In view of the fact that $\mathcal{I} \in \mathcal{F}$, it must be the case that $\mathcal{F} \cap \{x^*\}^\perp \subsetneq \mathcal{F}$. In either case we have $\mathcal{F}_{\min}^D \subsetneq \mathcal{F}$.

Reciprocally, if $\mathcal{F}_{\min}^D \subsetneq \mathcal{F}$, item ii. of Proposition 3 shows that there exists some $x \in \mathcal{F}^* \cap \ker \mathcal{A}$ such that either: i. $\langle c, x \rangle = 0$ and $x \notin \mathcal{F}^\perp$ or ii. $\langle c, x \rangle < 0$. Suppose first that i. holds, then the condition $x \notin \mathcal{F}^\perp$ readily implies that $\langle \mathcal{I}, x \rangle > 0$, by item v. of Lemma 2. So let $\alpha = \frac{1}{\langle \mathcal{I}, x \rangle}$. Then $(x\alpha, 0, 0)$ is an optimal solution for $(P_{\mathcal{F}})$, which shows that the optimal value is zero. Now suppose that ii. holds. We take $\alpha = \frac{1}{\langle \mathcal{I}, x \rangle - \langle c, x \rangle}$ and this is well-defined because $-\langle c, x \rangle > 0$. Then $(x\alpha, 0, -\alpha \langle c, x \rangle)$ is an optimal solution for $(P_{\mathcal{F}})$, which also shows that the optimal value is zero

- iii. If the common optimal value is nonzero, we must have $y_2^* > 0$ and $\mathcal{I}y_2^* \in \text{ri } \mathcal{F}$. This fact, together with Equation (12) and item iii. of Lemma 2, implies that $cy_1^* - \mathcal{A}^T y_3^* \in \text{ri } \mathcal{F}$ as well. Finally, $y_1^* \geq y_2^* > 0$, by Equation (13). Using Proposition 6, we can then conclude that $c - \mathcal{A}^T \frac{y_3^*}{y_1^*} \in \text{ri } \mathcal{F}_D^S$ as claimed. \square

Lemma 10 shows how to implement one step of a facial reduction algorithm using (IPO). In case the optimal value of $(P_{\mathcal{F}})$ is zero, we have two scenarios. In the first, we have $\langle c, x^* \rangle < 0$, we can then stop the procedure and declare that (D) is infeasible. In the second, the dual feasible region is contained in the face $\hat{\mathcal{F}} = \mathcal{F} \cap \{x^*\}^\perp$ and we have $\hat{\mathcal{F}} \subsetneq \mathcal{F}$. By Proposition 7, $\hat{\mathcal{F}}$ is linearly isomorphic to a smaller dimensional semidefinite cone \mathcal{S}_+^r with $r < n$, since $\hat{\mathcal{F}}$ is proper.

We may then reformulate $(\hat{\mathcal{F}}^*, \mathcal{A}, b, c)$ as a problem in a lower-dimensional space and apply Lemma 10 again. In the Appendix A, we will give a few more details about the technical issues regarding reformulations of the problem. Either way, we will stop either at the minimal face \mathcal{F}_{\min}^D or find out that the \mathcal{F}_D^S is actually empty. When facial reduction finally stops, we have one more pleasant surprise: item iii. of Proposition 7 shows that we can extract a relative interior solution from the dual problem $(D_{\mathcal{F}})$.

The upshot of this section is that we may use (IPO) to find \mathcal{F}_{\min}^D and if $\mathcal{F}_{\min}^D \neq \emptyset$ we can also obtain $s \in \text{ri } \mathcal{F}_D^S$. We write below a simplified facial reduction procedure for SDPs.

[Facial Reduction]

Input: $(\mathcal{S}_+^n, \mathcal{A}, b, c)$

Output: \mathcal{F}_{\min}^D and $s \in \text{ri } \mathcal{F}_D^S$ (if $\mathcal{F}_{\min}^D \neq \emptyset$), or a certificate of infeasibility (if $\mathcal{F}_{\min}^D = \emptyset$)

1) $\mathcal{F} \leftarrow \mathcal{S}_+^n$

- 2) Solve $(P_{\mathcal{F}})$ and $(D_{\mathcal{F}})$ using (IPO), to obtain primal dual pairs of optimal solutions (x^*, t^*, w^*) and (y_1^*, y_2^*, y_3^*) .
- 3) If $t^* = 0$ and $\langle c, x^* \rangle < 0$, let $\mathcal{F}_{\min}^D \leftarrow \emptyset$ and stop. $(\mathcal{S}_+^n, \mathcal{A}, b, c)$ is dual infeasible.
- 4) If $t^* = 0$ and $\langle c, x^* \rangle = 0$, let $\mathcal{F} \leftarrow \mathcal{F} \cap \{x^*\}^\perp$ and return to 2).
- 5) If $t^* > 0$, let $\mathcal{F}_{\min}^D \leftarrow \mathcal{F}$, $s \leftarrow c - \mathcal{A}^T \frac{y_1^*}{y_1}$ and stop.

Note that, if we denote by x^i the reducing direction of the i -th iteration of the facial reduction procedure, then the set of the reducing directions $\{x^1, \dots, x^\ell\}$ can be used as a finite certificate of infeasibility. The last direction x^ℓ will also satisfy (7) with \mathcal{K} possibly smaller than \mathcal{S}_+^n thus showing that $\mathcal{S}_+^n \cap \{x^1, \dots, x^{\ell-1}\}^\perp$ and $c - \text{range } \mathcal{A}^T$ can be strongly separated. Note that this does not imply that (D) is strongly infeasible, except if $\ell = 1$. To distinguish strong and weak infeasibility, it suffices to reformulate (7) with $\mathcal{K} = \mathcal{S}_+^n$ as a dual problem and apply the facial reduction procedure one more time to check its feasibility.

4 Hyper feasible partitions

Before we proceed, we need to introduce additional notation and terminology. Let $x \in \mathcal{S}^n$ be a symmetric matrix. For every $r \leq n$, we define the map $\pi_r : \mathcal{S}^n \rightarrow \mathcal{S}^m$ that takes x and maps to the upper left $r \times r$ principal submatrix of x . We will also define $\bar{\pi}_r : \mathcal{S}^n \rightarrow \mathcal{S}^{n-m}$, which maps x to the $(n-r) \times (n-r)$ lower right principal submatrix x . By convention, we have that $\pi_0 = \bar{\pi}_n$ is a map from \mathcal{S}^n to $\{0\} \subseteq \mathbb{R}$. Given a feasibility problem $(\mathcal{S}_+^n, \mathcal{L}, c)$ where \mathcal{L} is a subspace of \mathcal{S}^n , we will use $\pi_r(\mathcal{S}_+^n, \mathcal{L}, c)$ and $\bar{\pi}_r(\mathcal{S}_+^n, \mathcal{L}, c)$ as shorthand for $(\mathcal{S}_+^r, \pi_r(\mathcal{L}), \pi_r(c))$ and $(\mathcal{S}_+^{n-r}, \bar{\pi}_r(\mathcal{L}), \bar{\pi}_r(c))$, respectively. In the typical case, \mathcal{L} will be the image of \mathcal{A}^T . Finally, we say that $(\mathcal{S}_+^n, \mathcal{L}, c)$ is in *weak status* if it is either weakly feasible or weakly infeasible ([12]).

Suppose that there exists a nonzero element $a \in \mathcal{L} \cap \mathcal{S}_+^n$ with rank $r > 0$. Without loss of generality, we may assume that $\pi_r(a)$ is positive definite. If that is not the case, we may “rotate” the feasibility problem and consider the problem $(\mathcal{S}_+^n, u^T \mathcal{L} u, u^T c u)$, where u is an orthogonal matrix such that $\pi_r(u^T a u)$ is positive definite. We remark that u can be found by first computing an orthogonal basis for the kernel of a and extending it to an orthogonal basis of the whole space. This task can be accomplished, for instance, by making use of the Gram-Schmidt process. Note that only elementary real arithmetic operations and square roots are needed for that and we already showed that square roots can be computed using (IPO). In Theorem 12 of [12], the authors proved the following result.

Theorem 11. *Let $(\mathcal{S}_+^n, \mathcal{L}, c)$ be a semidefinite feasibility problem and suppose that there is a rank $r > 0$ matrix $a \in \mathcal{L} \cap \mathcal{S}_+^n$ such that $\pi_r(a)$ is positive definite. Then*

1. $(\mathcal{S}_+^n, \mathcal{L}, c)$ is strongly feasible if and only if $\bar{\pi}_r(\mathcal{S}_+^n, \mathcal{L}, c)$ is.
2. $(\mathcal{S}_+^n, \mathcal{L}, c)$ is strongly infeasible if and only if $\bar{\pi}_r(\mathcal{S}_+^n, \mathcal{L}, c)$ is.
3. $(\mathcal{S}_+^n, \mathcal{L}, c)$ is in weak status if and only if $\bar{\pi}_r(\mathcal{S}_+^n, \mathcal{L}, c)$ is.

The idea is that as long as there is a matrix $a \in \mathcal{S}_+^{n-r} \cap \bar{\pi}_r(\mathcal{L})$ we may, rotating the problem if necessary, apply Theorem 11 to obtain smaller problems having almost the same feasibility status. As before, only elementary arithmetic operations and square roots are needed for rotating the problem. The following result was proved in [12].

Theorem 12. *Suppose that \mathcal{L} is a subspace such that there is a nonzero element in $\mathcal{S}_+^n \cap \mathcal{L}$. Then, there is an orthogonal matrix P and $\{A_1, \dots, A_\ell\} \subseteq P^T \mathcal{L} P$ such that the following hold.*

- i. $\ell \leq n - 1$.
- ii. $A_1 \in \mathcal{S}_+^n$, $\pi_{k_1}(A_1)$ is positive definite and $\bar{\pi}_{k_1}(A_1) = 0$, where $k_1 = \text{rank } A_1$.

iii. For every $i > 1$, A_i is such that $\pi_{k_i}(\bar{\pi}_{k_1+\dots+k_{i-1}}(A_i))$ is positive definite and $\bar{\pi}_{k_1+\dots+k_i}(A_i) = 0$, where $k_i = \text{rank } \bar{\pi}_{k_1+\dots+k_{i-1}}(A_i)$.

iv. There is no $A \in P^T \mathcal{L} P$ such that $\bar{\pi}_{k_1+\dots+k_\ell}(A)$ is positive semidefinite and nonzero.

Moreover, for every $c \in \mathcal{S}^n$ and non-singular matrix P , and for any subset $\{A_1, \dots, A_\ell\} \subseteq P^T \mathcal{L} P$ satisfying items ii. through iv, the following relations holds between $(\mathcal{S}_+^n, \mathcal{L}, c)$ and $\bar{\pi}_{k_1+\dots+k_\ell}(\mathcal{S}_+^n, P^T \mathcal{L} P, P^T c P)$:

1. $(\mathcal{S}_+^n, \mathcal{L}, c)$ is strongly feasible if and only if $\bar{\pi}_{k_1+\dots+k_\ell}(\mathcal{S}_+^n, P^T \mathcal{L} P, P^T c P)$ is.
2. $(\mathcal{S}_+^n, \mathcal{L}, c)$ is strongly infeasible if and only if $\bar{\pi}_{k_1+\dots+k_\ell}(\mathcal{S}_+^n, P^T \mathcal{L} P, P^T c P)$ is.
3. $(\mathcal{S}_+^n, \mathcal{L}, c)$ is in weak status if and only if $\bar{\pi}_{k_1+\dots+k_\ell}(\mathcal{S}_+^n, P^T \mathcal{L} P, P^T c P)$ is weakly feasible.

This result summarizes part of discussion presented in [12]. Proof of existence and of the properties of items 1) to 3) are contained in Proposition 19 and in the description of the Algorithm 1 therein. The bound on ℓ comes from Theorem 23 in [12]. We remark that the notation is slightly different, here we use \mathcal{S}_+^n , π_k and $\bar{\pi}_k$ instead of K_n , U_k and L_k .

The $\{A_1, \dots, A_\ell\}$ together with $\{k_1, \dots, k_\ell\}$ form a so-called *maximal hyper feasible partition* of \mathcal{L} . The matrices are such that

$$\begin{pmatrix} \hat{A}_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} * & * & * & * \\ * & \hat{A}_2 & 0 & 0 \\ * & 0 & 0 & 0 \\ * & 0 & 0 & 0 \end{pmatrix}, \dots, \begin{pmatrix} * & * & * & * \\ * & * & * & * \\ * & * & \hat{A}_\ell & 0 \\ * & * & 0 & 0 \end{pmatrix},$$

where each \hat{A}_i is a $k_i \times k_i$ positive definite matrix. The term “maximal” is used because of item iv. of Theorem 12. Also in [12], there is a description of a procedure to compute a hyper feasible partition, see Algorithm 1 therein. Its core is finding a nonzero element in $\mathcal{S}_+^n \cap \mathcal{L}$ or checking that none exists. Everything else is just to keep track of the subproblems and rotate the problems and the matrices in appropriate manner. In the next proposition, we discuss how (IPO) can be used to find a nonzero element in $\mathcal{S}_+^n \cap \mathcal{L}$. This will show that a maximal hyper feasible partition can be found using (IPO). In what follows, we let $\mathcal{I} \in \mathcal{S}^n$ denote the identity matrix.

Proposition 13. Let \mathcal{L} be a subspace of \mathcal{S}^n and let \mathcal{B} be a linear map and d a vector such that the following equality holds.

$$\begin{aligned} & \{(s, t) \in \mathcal{S}^n \times \mathbb{R} \mid \mathcal{B}(s, t) = d\} \\ & = \{(s, t) \in \mathcal{S}^n \times \mathbb{R} \mid \exists l \in \mathcal{L} \text{ s.t. } s = l + t\mathcal{I}, \langle \mathcal{I}, s \rangle = 1\} \end{aligned}$$

Now consider the following problem

$$\begin{aligned} & \underset{s, t}{\text{minimize}} \quad \langle (s, t), (0, 1) \rangle \\ & \text{subject to} \quad \mathcal{B}(s, t) = d \\ & \quad (s, t) \in \mathcal{S}_+^n \times \mathcal{S}_+^1 \end{aligned} \tag{14}$$

It has the following properties:

- i. Both (14) and its dual have interior feasible points (so we may invoke (IPO) to solve it).
- ii. Let (s^*, t^*) be an optimal solution. The optimal value is zero if and only if there is a nonzero element in $\mathcal{S}_+^n \cap \mathcal{L}$. In that case, s^* is a nonzero element in $\mathcal{S}_+^n \cap \mathcal{L}$.

Proof. i. Note that $(\frac{\mathcal{I}}{n}, 1)$ is an interior solution for (14). To prove that the dual problem has an interior solution, we use item ii. Proposition 3. If (s, t) satisfies $\mathcal{B}(s, t) = 0$, $t \leq 0$ and $(s, t) \in \mathcal{S}_+^n \times \mathcal{S}_+^1$ then $t = 0$ and $\langle \mathcal{I}, s \rangle = 0$, which forces $s = 0$.

ii. Let (s^*, t^*) be an optimal solution. If $t^* = 0$, then $s^* \in \mathcal{L}$ and s^* is nonzero since $\langle \mathcal{I}, s \rangle = 1$. Reciprocally, let s be a nonzero element in $\mathcal{S}_+^n \cap \mathcal{L}$, then $(\frac{s}{\langle \mathcal{I}, s \rangle}, 0)$ is an optimal solution for (14). \square

Note that the directions that compose a hyper feasible partition do not belong to \mathcal{L} but to $P^T \mathcal{L} P$, where P is some orthogonal matrix. However, we can “rotate” $(\mathcal{S}_+^n, \mathcal{A}, b, c)$ by P and consider the problem $(\mathcal{S}_+^n, \mathcal{B}, b, P^T c P)$ instead, where \mathcal{B} satisfies $\mathcal{B}^T y = \sum_{i=1}^m P^T \mathcal{A}_i P y$, for every y . $(\mathcal{S}_+^n, \mathcal{B}, b, P^T c P)$ is equivalent to $(\mathcal{S}_+^n, \mathcal{A}, b, c)$ and both share the same primal and dual optimal values.

Before we close this section, we will show that when (D) is not strongly infeasible, we can use (IPO) together with a maximal hyper feasible partition to obtain points in $c - \text{range } \mathcal{A}^T$ which are arbitrarily close to \mathcal{S}_+^n . We first need the following lemma.

Lemma 14. *Let $\{A_1, \dots, A_\ell\} \subseteq \mathcal{S}^n$ be a subset which satisfies items ii and iii. of Theorem 12. Let $s = k_1 + \dots + k_\ell$ and suppose that $z \in \mathcal{S}^n$ is such that $\bar{\pi}_s(z)$ is positive definite. Then, there is a linear combination of the A_i such that $z + \sum_{i=1}^\ell A_i \alpha_i$ is positive definite and all the α_i are positive.*

Proof. Let \mathcal{L} be the space spanned by the A_i . Due to the assumption on z , we have that $\bar{\pi}_s(\mathcal{S}_+^n, \mathcal{L}, z)$ is strongly feasible. Now, item iv of Theorem 12 implies that $(\mathcal{S}_+^n, \mathcal{L}, z)$ is strongly feasible as well, so $z + l$ is positive definite for some $l \in \mathcal{L}$.

The assertion about the positiveness of the coefficients α_i comes from a constructive way of building $z + l$. Since $\pi_{k_\ell}(\bar{\pi}_{k_1+\dots+k_{\ell-1}}(A_\ell))$ is positive definite and $\bar{\pi}_{k_1+\dots+k_\ell}(A_\ell) = 0$, we have that $\bar{\pi}_{k_1+\dots+k_{\ell-1}}(\alpha_\ell A_\ell + z)$ has the shape $\begin{pmatrix} \alpha_\ell F & * \\ * & H \end{pmatrix}$, where both $F \in \mathcal{S}_+^{k_\ell}$ and $H \in \mathcal{S}_+^{n-s}$ are positive definite matrices and $*$ denotes arbitrary entries. An argument with the Schur complement shows that if α_ℓ is positive and sufficiently large then $\bar{\pi}_{k_1+\dots+k_{\ell-1}}(\alpha_\ell A_\ell + z)$ is positive definite as well. By the same argument, there exists a positive $\alpha_{\ell-1}$ such that $\bar{\pi}_{k_1+\dots+k_{\ell-2}}(\alpha_{\ell-1} A_{\ell-1} + \alpha_\ell A_\ell + z)$ is positive definite. By induction, we can prove the assertion about the coefficients. See also Section 4.2 of [12], where a similar argument is used. \square

In what follows, let $e_j \in \mathbb{R}^n$ denote the vector that has 1 in the j -position and 0 elsewhere. Moreover, let $\mathcal{I} \in \mathcal{S}_+^n$ denote the identity matrix and $\mathbb{R}_+^\ell = \{x \in \mathbb{R}^n \mid x_i \geq 0, \forall i\}$.

Proposition 15. *Let $\{A_1, \dots, A_\ell\} \subseteq \mathcal{S}^n$ be a subset which satisfies items ii and iii. of Theorem 12. Let $s = k_1 + \dots + k_\ell$ and suppose that $z \in \mathcal{S}^n$ is such that $\bar{\pi}_s(z)$ is positive semidefinite.*

Let $\epsilon > 0$ and consider the following pair of primal and dual problems.

$$\begin{aligned} & \underset{\alpha}{\text{maximize}} && - \sum_{i=1}^\ell \alpha_i && (D_\epsilon) \\ & \text{subject to} && \left(z + \frac{\epsilon}{n} \mathcal{I}, 0 \right) - \sum_{i=1}^\ell (-A_i, -e_i) \alpha_i && \in \mathcal{S}_+^n \times \mathbb{R}_+^\ell, \end{aligned}$$

The corresponding primal is

$$\begin{aligned} & \underset{x}{\text{minimize}} && \left\langle \left(z + \frac{\epsilon}{n} \mathcal{I}, 0 \right), x \right\rangle && (P_\epsilon) \\ & \text{subject to} && \langle (-A_i, -e_i), x \rangle = -1, \quad i = 1, \dots, \ell && (15) \\ & && x \in \mathcal{S}_+^n \times \mathbb{R}_+^\ell. \end{aligned}$$

The following properties hold.

- i. Both (D_ϵ) and (P_ϵ) have interior feasible points.
- ii. Let α^* be an optimal solution to (D), then $\tilde{z} = z + \sum_{i=1}^\ell A_i \alpha_i^*$ is such that $\text{dist}(\tilde{z}, \mathcal{S}_+^n) \leq \epsilon$.

Proof. *i.* We first show that (D_ϵ) has an interior feasible point. Recall that $\bar{\pi}_s(z)$ is positive semidefinite. Since \mathcal{I} is positive definite, $\pi_s(\tilde{z} + \frac{\epsilon}{n}\mathcal{I})$ is positive definite as well and we may invoke Lemma 14 to obtain α in the interior of \mathbb{R}_+^ℓ such that $z + \epsilon\mathcal{I} - \sum_{i=1}^\ell A_i\alpha_i$ is positive definite. This shows that (D_ϵ) is strongly feasible.

We will now use item *i*. Proposition 3 to show that (P_ϵ) is strongly feasible. First, note that a solution to the linear system (15) is given by $(0, e_1 + \dots + e_\ell)$. Moreover, every α satisfying both $\sum_{i=1}^\ell (-A_i, -e_i)\alpha_i \in \mathcal{S}_+^n \times \mathbb{R}_+^\ell$ and $-\sum_{i=1}^\ell -\alpha_i \geq 0$ must be zero.

ii. Since $\tilde{z} + \epsilon\mathcal{I}$ is positive semidefinite, we have that $\text{dist}(\tilde{z}, \mathcal{S}_+^n) \leq \text{dist}(\tilde{z}, \tilde{z} + \frac{\epsilon}{n}\mathcal{I}) = \sqrt{\text{trace}(\frac{\epsilon^2}{n^2}\mathcal{I})} = \epsilon$. \square

Corollary 16. *Under the conditions of Proposition 15, if in addition we have that $\bar{\pi}_s(z)$ is positive definite, the result also holds with $\epsilon = 0$. So, in particular, $\tilde{z} = z + \sum_{i=1}^\ell A_i\alpha_i^* \in \mathcal{S}_+^n$.*

Proof. Existence of an interior feasible for (D_ϵ) is guaranteed by Lemma 14, one then proceeds exactly as in the proof of Proposition 15. \square

Suppose that (D) is not strongly infeasible and let $\{D_1, \dots, D_\ell\}$ be a maximal hyper feasible partition for $\text{range } \mathcal{A}^T$ and assume, if necessary, that the problem was already rotated by an appropriate orthogonal matrix so that the $D_i \in \text{range } \mathcal{A}^T$, for all i . Then the last problem $\bar{\pi}_s(\mathcal{S}_+^n, \mathcal{L}, c)$ is feasible by Theorem 12, where $s = k_1 + \dots + k_\ell$. Let \hat{y} be any feasible solution for $\bar{\pi}_s(\mathcal{S}_+^n, \mathcal{L}, c)$. Recall that \hat{y} can be obtained via facial reduction as in Section 3. Then we can use Proposition 15 to show that for every $\epsilon > 0$ there is a non-negative linear combination of the D_i for which $\text{dist}(c - \mathcal{A}^T\hat{y} + \sum_{i=1}^\ell \alpha_i D_i, \mathcal{S}_+^n) \leq \epsilon$ and that we may find the α_i using (IPO). In particular, this also applies to weakly infeasible problems so that the hyper feasible partition can be used to produce points that are arbitrarily close to feasibility.

5 Finding the optimal value

The current status is that we already know we may find the \mathcal{F}_{\min}^D with finitely many calls to (IPO). However, if we compute the primal problem with respect to \mathcal{F}_{\min}^D there is no guarantee that it will be strongly feasible as well. Nevertheless, we will show how we can find the optimal value of (D) using (IPO).

Our assumption here is that \mathcal{F}_{\min}^D was found and the problem was reformulated into a dual standard form problem (D) which is strongly feasible. We shall consider two cases, $b = 0$ and $b \neq 0$. The former can be treated easily, since the optimal value is zero. For the latter, we may assume without loss of generality that $b_1 = 1$ and define $y_0 = y_1 + \dots + b_m y_m$. Replacing each \mathcal{A}_i for $\mathcal{A}_i - b_i \mathcal{A}_1$, we can transform (D) into the following equivalent problem:

$$\begin{aligned} & \underset{y=(y_0, y_2, \dots, y_m)}{\text{maximize}} && y_0 && (D_{\text{ref}}) \\ & \text{subject to} && c - \mathcal{A}_1 y_0 - \sum_{i=2}^m \mathcal{A}_i y_i \in \mathcal{S}_+^n. \end{aligned}$$

Theorem 17. *Suppose that (D_{ref}) is strongly feasible. and let \mathcal{L} be the span of $\{\mathcal{A}_2, \dots, \mathcal{A}_m\}$. Suppose that there is a maximal hyper feasible partition $\{D_1, \dots, D_\ell\}$ for \mathcal{L} and k_1, \dots, k_ℓ be as in Theorem 11. In this case, define $s = k_1 + \dots + k_\ell$ and, if necessary, rotate the problem so that all the $D_i \in \text{range } \mathcal{A}^T$. If there is no maximal hyper feasible partition, then let $s = 0$ ¹. The following problem has the same optimal value as*

¹This can only happen if $\mathcal{L} \cap \mathcal{S}_+^n = \{0\}$.

(D_{ref})

$$\begin{aligned} & \underset{y}{\text{maximize}} \quad y_0 & (D_{\text{small}}) \\ & \text{subject to} \quad \bar{\pi}_s \left(c - \mathcal{A}_1 y_0 - \sum_{i=2}^m \mathcal{A}_i y_i \right) \in \mathcal{S}_+^{n-s}. \end{aligned}$$

The corresponding primal is

$$\underset{x}{\text{minimize}} \quad \langle \bar{\pi}_s(c), x \rangle \quad (P_{\text{small}})$$

$$\text{subject to} \quad \langle \bar{\pi}_s(\mathcal{A}_1), x \rangle = 1 \quad (16)$$

$$\begin{aligned} & \langle \bar{\pi}_s(\mathcal{A}_j), x \rangle = 0, \quad j = 2, \dots, m \\ & x \in \mathcal{S}_+^{n-s}. \end{aligned} \quad (17)$$

There are two possibilities for (P_{small}) :

1. (P_{small}) is strongly infeasible and $\theta_{D_{\text{ref}}} = +\infty$.
2. (P_{small}) has an interior feasible solution.

Proof. For the proof that $\theta_{D_{\text{small}}} = \theta_{D_{\text{ref}}}$, we only need to consider the case $s > 0$. The first observation is that $\theta_{D_{\text{small}}} \geq \theta_{D_{\text{ref}}}$, since the feasible region of (D_{small}) contains the feasible region of (D_{ref}) . We will now show that $\theta_{D_{\text{small}}} \leq \theta_{D_{\text{ref}}}$ holds as well. Due to Theorem 11, (D_{small}) must be strongly feasible. This means that, for all $\mu < \theta_{D_{\text{small}}}$ there is $(y_0^*, y_2^*, \dots, y_m^*)$ such that $\pi_s(c - \mathcal{A}_1 y_0^* - \sum_{i=2}^m \mathcal{A}_i y_i^*)$ is positive definite and $y_0^* \geq \mu$. However, there is a linear combination of the directions $\{D_1, \dots, D_r\}$ such that $z = c - \mathcal{A}_1 y_0^* - \sum_{i=2}^m \mathcal{A}_i y_i^* + \sum_{j=1}^r \alpha_j D_j$ is positive definite, by Lemma 14. As all the D_j lie in the span of $\{\mathcal{A}_2, \dots, \mathcal{A}_m\}$ we have that z corresponds to a feasible solution \hat{y} such that $\hat{y}_0 = y_0^*$. This shows that $\theta_{D_{\text{small}}} = \theta_{D_{\text{ref}}}$.

Now we prove the statement about the primal problem (P_{small}) . First, suppose that the linear system composed of (16) and (17) is infeasible. From basic linear algebra, we have that $(1, 0, \dots, 0)$ can be written as a sum $u + v$, with v satisfying $\bar{\pi}_s(\mathcal{A}^T v) = 0$ and $\langle u, v \rangle = 0$. Therefore, if y is any feasible solution to (D_{small}) , then $y + \alpha v$ will also be a feasible solution for any $\alpha \in \mathbb{R}$. This shows that $\theta_{D_{\text{small}}} = +\infty$.

Finally, suppose that (16) and (17) is feasible and that (P_{small}) does not have an interior solution. By item *i.* of Proposition 3, there is some $y = (y_0, y_2, \dots, y_m)$ such that $y_0 \geq 0$ and $z = -\bar{\pi}_s(\mathcal{A}_1 y_0 + \sum_{i=2}^m \mathcal{A}_i y_i)$ is a nonzero positive semidefinite matrix. If $y_0 = 0$ and $s > 0$, we have that $-\bar{\pi}_s(\sum_{i=2}^m \mathcal{A}_i y_i)$ is a positive semidefinite matrix which must be zero, due to item *iv.* of Theorem 11. If $s = 0$ and $y_0 = 0$, we have $\mathcal{L} \cap \mathcal{S}_+^n = \{0\}$, which also implies that z is zero. Both possibilities contradict the fact that z is nonzero, so we must have $y_0 > 0$ and z must be a certificate of strong infeasibility for the primal problem, by Proposition 4. In this case, Corollary 5 implies that $\theta_{D_{\text{ref}}} = +\infty$. □

Example 18. Consider the following SDP in dual format.

$$\begin{aligned} & \underset{y_1, y_2, y_3}{\text{maximize}} \quad -y_1 - y_2 - y_3 & (D_{\text{ex}}) \\ & \text{subject to} \quad \begin{pmatrix} y_1 & 1 & y_2 \\ 1 & y_2 & 1 \\ y_2 & 1 & y_1 + y_2 + y_3 \end{pmatrix} \in \mathcal{S}_+^3, \end{aligned}$$

Note that (D_{ex}) is strongly feasible, so we can apply the techniques described in this section. We consider the change of variables $y_0 = -y_1 - y_2 - y_3$ and write down the following equivalent problem.

$$\begin{aligned} & \underset{y_0, y_1, y_2}{\text{maximize}} \quad y_0 & (D_{\text{ex}2}) \\ & \text{subject to} \quad \begin{pmatrix} y_1 & 1 & y_2 \\ 1 & y_2 & 1 \\ y_2 & 1 & -y_0 \end{pmatrix} \in \mathcal{S}_+^3. \end{aligned}$$

The next step is to compute a maximal hyper feasible partition for the subspace

$$\mathcal{L} = \left\{ \begin{pmatrix} y_1 & 0 & y_2 \\ 0 & y_2 & 0 \\ y_2 & 0 & 0 \end{pmatrix} \mid y_1, y_2 \in \mathbb{R} \right\}.$$

We can take, for instance, $D_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ and $D_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$. We then have $k_1 = k_2 = 1$. Following Theorem 17, the next SDP has the same optimal value as (D_{ex2}) .

$$\begin{aligned} & \underset{y_0}{\text{maximize}} && y_0 && (D_{\text{ex3}}) \\ & \text{subject to} && (-y_0) \in \mathcal{S}_+^1. \end{aligned}$$

The optimal value of (D_{ex3}) is zero and is attained, even though the optimal value of (D_{ex2}) is unattained.

Theorem 17 implies that in order to compute the optimal value θ_D , one must first check whether (P_{small}) is strongly infeasible or not. By Proposition 4, strong infeasibility can be detected by solving a SDP feasibility problem and the discussion in Section 3 shows how to do that with facial reduction and (IPO). If (P_{small}) is not strongly infeasible, then (IPO) can be invoked using (P_{small}) and (D_{small}) as input.

The approach described in this section can be seen as trying to eliminate common recession directions between the objective function and the feasible set, this line of research has been pursued in the past by Abrams in [1].

6 Recovering optimal solution

Here we suppose that θ_D is known and it is finite. This means that for every $\epsilon > 0$ there exists y such that $c - \mathcal{A}^T y \in \mathcal{S}_+^n$ and $\langle b, y \rangle \geq \theta_D - \epsilon$. It follows that $\{s \in \mathcal{S}^n \mid \exists y \text{ s.t. } s = c - \mathcal{A}^T y, \langle b, y \rangle = \theta_D\}$ is a non-empty affine space, which can be written in the form $\widehat{\mathcal{L}} + \widehat{c}$, for some $\widehat{c} \in \mathcal{S}^n$ and some subspace $\widehat{\mathcal{L}} \subseteq \mathcal{S}^n$. Note that the feasibility problem $(\mathcal{S}_+^n, \widehat{\mathcal{L}}, \widehat{c})$ corresponds to the optimal set of (D) and it cannot be strongly infeasible. After all, there are points in \mathcal{S}_+^n which either satisfy or almost satisfy the linear equalities which define $\widehat{\mathcal{L}} + \widehat{c}$, by the definition of θ_D . Since $\widehat{\mathcal{L}} + \widehat{c}$ is polyhedral, this fact is enough to establish $\text{dist}(\mathcal{S}_+^n, \widehat{\mathcal{L}} + \widehat{c}) = 0$.

When the dual optimal value is not attained, it must be the case that $(\mathcal{S}_+^n, \widehat{\mathcal{L}}, \widehat{c})$ is weakly infeasible. Infeasibility of $(\mathcal{S}_+^n, \widehat{\mathcal{L}}, \widehat{c})$ can, once again, be detected via the facial reduction procedure of Section 3. Note that if the problem turns out to be feasible, facial reduction will also return a point that belongs to the relative interior of the optimal set. Due to the facial structure of \mathcal{S}_+^n this corresponds to an optimal solution of maximal rank.

For the unattained case, the discussion in Sections 4 and 5 can be used to generate feasible points which are arbitrarily close to optimality. Under the setting of Theorem 17, the problem (D_{small}) is always attained, even if (D) is not. So let y^* be an optimal solution to (D_{small}) . Also, let \hat{y} be any vector for which $\bar{\pi}_s(c - \mathcal{A}_1 \hat{y}_0 - \sum_{i=2}^m \mathcal{A}_i \hat{y}_i)$ is an interior point of \mathcal{S}_+^{n-s} . Note that \hat{y} is obtainable through a single facial reduction step by item *iii.* of Lemma 10, since (D_{small}) is strongly feasible. For every $\gamma \in [0, 1)$, consider the matrix

$$z_\gamma = \bar{\pi}_s \left(c - \mathcal{A}_1((1 - \gamma)\hat{y}_0 + \gamma y_0^*) - \sum_{i=2}^m \mathcal{A}_i((1 - \gamma)\hat{y}_i + \gamma y_i^*) \right).$$

By the choice of \hat{y} , we have that z_γ is a feasible slack for (D_{small}) and it is positive definite for all $\gamma \in [0, 1)$. We can then apply Corollary 16 and use (IPO) to obtain a feasible slack for (D_{ref}) in the format

$$\tilde{z}_\gamma = c - \mathcal{A}_1((1 - \gamma)\hat{y}_0 + \gamma y_0^*) - \sum_{i=2}^m \mathcal{A}_i((1 - \gamma)\hat{y}_i + \gamma y_i^*) + \sum_{i=1}^\ell \alpha_i D_i,$$

where the coefficients α_i depend on the choice of γ . Adding elements of the hyper feasible partition do not affect the value of the solutions, so \tilde{z}_γ corresponds to some feasible solution to (D_{ref}) that has value equal

to $(1 - \gamma)\hat{y}_0 + \gamma y_0^*$, where y_0^* is the optimal value. Since the problem is unattained, we have that $\hat{y}_0 < y_0^*$. Therefore, if we wish a feasible point to (D_{ref}) whose optimal value is within some $\epsilon > 0$ of the true optimal, it is enough to take $\gamma > \frac{y_0^* - \hat{y}_0 - \epsilon}{y_0^* - \hat{y}_0}$.

7 Gluing everything together

Here, we summarize the steps that one may take to completely solve (D) using the (IPO).

- 1) Use facial reduction to find \mathcal{F}_{\min}^D . If $\mathcal{F}_{\min}^D = \emptyset$, then (D) is infeasible.
 - (a) If (D) is infeasible, solve the feasibility problem in item *ii* of Proposition 4, using facial reduction. If that problem is infeasible, then (D) is weakly infeasible. If it is feasible, then (D) is strongly infeasible.
 - (b) If (D) turns out to be weakly infeasible, there will be a maximal hyper feasible partition, see Proposition 4 in [12]. In this case, the discussion at the end of Section 4 shows how to produce points arbitrarily close to the positive semidefinite cone.
- 2) Reformulate $(\mathcal{S}_+^n, \mathcal{A}, b, c)$ as a dual strongly feasible SDP over the face \mathcal{F}_{\min}^D . See Appendix A for details.
- 3) Compute a hyper feasible partition for the reformulated problem and proceed as in Section 5. At this point, the optimal value θ_D will be known. If $\theta_D = +\infty$, we stop and declare that the dual is unbounded. This happens if and only if (P_{small}) is strongly infeasible.
- 4) Knowing that $\theta_D < +\infty$, we proceed as in Section 6 and formulate the feasibility problem and apply Facial Reduction to it, as in Section 3.
 - (a) If $(\mathcal{S}_+^n, \hat{\mathcal{L}}, \hat{c})$ is feasible, we obtain an optimal solution of maximal rank as a byproduct of facial reduction.
 - (b) If $(\mathcal{S}_+^n, \hat{\mathcal{L}}, \hat{c})$ is (weakly) infeasible, we can proceed as in Section 6 to obtain feasible points arbitrarily close to optimality.

8 Further discussion

We note that de Klerk, Terlaky and Roos have described in section 5.10 of [7], a possible sequence of steps to solve (D). Their tool of choice is a self-dual embedding strategy of the original pair (P) and (D). As we mentioned before, in the absence of both primal and dual strong feasibility, the embedded problem might fail to reveal the optimal value of the original problem or detect infeasibility/nonattainment. To account for that, they go for a second step, where they consider an embedded problem using Ramana's dual. The Ramana's dual (P_R) is a substitute for (P) and they consider the pair formed by (P_R) and its dual (D_{cor}) , which is a "corrected" version of (D). The pair (P_R, D_{cor}) can then be solved by their embedding strategy to find θ_D . As the embedded problem is both primal and dual strongly feasible, it is possible to invoke (IPO) to solve it. However, if the solution given by (IPO) is not of maximum rank at both steps, their strategy might not work. We should mention that they do show in detail how to build an interior point method suitable for their approach. Our analysis, on the other hand, is completely agnostic to the inner workings of the interior point oracle and no assumption is made on the optimal solutions returned by (IPO).

Nevertheless, missing from their analysis is how one can recover a solution to (D) from a solution (D_{cor}) or to check that this is not possible. Moreover, it is not clear from their approach how to obtain points close to optimality in case of unattainment, or points close to feasibility in case of weak infeasibility. As our approach does not rely on Ramana's dual, we believe our analysis is more easily generalizable to other classes of cones. We remark that although there is a strong connection between Ramana's dual and facial reduction [24, 19], no similar construction is known for any other class of cones. For example, following Pataki's approach in [19], one could formulate an alternative dual system for a second order cone programming problem. Such a

system would have many of the properties that Ramana’s dual has, but it is not clear whether that system can be expressed via second order cone constraints. That is, we may fall outside the problem class under consideration.

9 Concluding remarks

Many of the results described here are already in a form suitable for further generalization and we intend to explore this theme in a future work. For instance the pair of problems $(P_{\mathcal{F}})$ and $(D_{\mathcal{F}})$ can be used to carry out facial reduction for arbitrary closed convex cones and they are strongly feasible regardless of the choice of \mathcal{K} . On the other hand, since the positive semidefinite cone has a very nice structure, it was possible to simplify the discussion of subtle but important technical points. For example, we mentioned in Section 3 that, in principle, we might have needed a stronger oracle that is able to solve problems over proper faces of \mathcal{K} as well. Since proper faces of \mathcal{S}_+^n are linearly isomorphic to smaller SDPs, no such assumption was necessary. It seems that for general cones we might need to resort to this type of assumption.

Acknowledgements

We thank Henry Wolkowicz for bringing to our attention the paper by Abrams [1]. The first author would also like to acknowledge the support and the interesting discussions with Prof. Mitsuhiro Fukuda at the Tokyo Institute of Technology.

References

- [1] Robert A. Abrams. Projections of convex programs with unattained infima. *SIAM Journal on Control*, 13(3):706–718, 1975.
- [2] F. Alizadeh. Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM J. Optim.*, 5(1):13–51, 1995.
- [3] Lenore Blum, Felipe Cucker, Michael Shub, and Steve Smale. *Complexity and Real Computation*. Springer, New York, NY, October 1998.
- [4] Jonathan M. Borwein and Henry Wolkowicz. Facial reduction for a cone-convex programming problem. *Journal of the Australian Mathematical Society (Series A)*, 30(03):369–380, 1981.
- [5] Jonathan M. Borwein and Henry Wolkowicz. Regularizing the abstract convex program. *Journal of Mathematical Analysis and Applications*, 83(2):495 – 530, 1981.
- [6] Yuen-Lam Cheung, Simon Schurr, and Henry Wolkowicz. Preprocessing and regularization for degenerate semidefinite programs. In *Computational and Analytical Mathematics*, volume 50 of *Springer Proceedings in Mathematics & Statistics*, pages 251–303. Springer New York, 2013.
- [7] Etienne de Klerk, Tamás Terlaky, and Kees Roos. Self-dual embeddings. In Henry Wolkowicz, Romesh Saigal, and Lieven Vandenbergh, editors, *Handbook of semidefinite programming: theory, algorithms, and applications*, chapter 5. Kluwer Academic Publishers, 2000.
- [8] Katsuki Fujisawa, Mitsuhiro Fukuda, Kazuhiro Kobayashi, Masakazu Kojima, Kazuhide Nakata, Maho Nakata, and Makoto Yamashita. SDPA (SemiDefinite Programming Algorithm) and SDPA-GMP User’s Manual — version 7.1.1. Technical Report B-448, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, 2008.
- [9] Igor Klep and Markus Schweighofer. An exact duality theory for semidefinite programming based on sums of squares. *Math. Oper. Res.*, 38(3):569–590, August 2013.

- [10] Minghui Liu and Gábor Pataki. Exact duals and short certificates of infeasibility and weak infeasibility in conic linear programming. Technical report, Department of Statistics and Operations Research, University of North Carolina at Chapel Hill, 2015.
- [11] Minghui Liu and Gábor Pataki. Exact duality in semidefinite programming based on elementary reformulations. *SIAM Journal on Optimization*, 25(3):1441–1454, 2015.
- [12] Bruno F. Lourenço, Masakazu Muramatsu, and Takashi Tsuchiya. A structural geometrical analysis of weakly infeasible SDPs. Technical report, 2013.
- [13] Z. Luo, J. F. Sturm, and S. Zhang. Duality results for conic convex programming. Technical report, Econometric Institute, Erasmus University Rotterdam, The Netherlands, 1997.
- [14] Z. Luo, J. F. Sturm, and S. Zhang. Conic convex programming and self-dual embedding. *Optimization Methods and Software*, 14(3):169–218, 2000.
- [15] Y. Nesterov and A. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. Society for Industrial and Applied Mathematics, 1994.
- [16] Y. Nesterov, M. J. Todd, and Y. Ye. Infeasible-start primal-dual methods and infeasibility detectors for nonlinear programming problems. *Mathematical Programming*, 84(2):227–267, February 1999.
- [17] Gábor Pataki. The geometry of semidefinite programming. In Henry Wolkowicz, Romesh Saigal, and Lieven Vandenbergh, editors, *Handbook of semidefinite programming: theory, algorithms, and applications*. Kluwer Academic Publishers, 2000.
- [18] Gábor Pataki. Bad semidefinite programs: they all look the same. Technical report, Department of Statistics and Operations Research, University of North Carolina at Chapel Hill, 2014.
- [19] Gábor Pataki. Strong duality in conic linear programming: Facial reduction and extended duals. In *Computational and Analytical Mathematics*, volume 50, pages 613–634. Springer New York, 2013.
- [20] F. Permenter and P. Parrilo. Partial facial reduction: simplified, equivalent SDPs via approximations of the PSD cone. *ArXiv e-prints*, 2014.
- [21] Imre Pólik and Tamás Terlaky. New stopping criteria for detecting infeasibility in conic optimization. *Optimization Letters*, 3(2):187–198, March 2009.
- [22] Florian A. Potra and Rongqin Sheng. On homogeneous interior-point algorithms for semidefinite programming. *Optimization Methods and Software*, 9(1-3):161–184, 1998.
- [23] Motakuri V. Ramana. An exact duality theory for semidefinite programming and its complexity implications. *Mathematical Programming*, 77, 1995.
- [24] Motakuri V. Ramana, Levent Tunçel, and Henry Wolkowicz. Strong duality for semidefinite programming. *SIAM Journal on Optimization*, 7(3):641–662, August 1997.
- [25] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, 1997.
- [26] Jos F. Sturm. Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11(1-4):625–653, 1999.
- [27] Jos F. Sturm. Error bounds for linear matrix inequalities. *SIAM Journal on Optimization*, 10(4):1228–1248, January 2000.
- [28] Kim-Chuan Toh, Michael J. Todd, and Reha H. Tütüncü. On the implementation and usage of SDPT3 – a Matlab software package for semidefinite-quadratic-linear programming, version 4.0. In Miguel F. Anjos and Jean B. Lasserre, editors, *Handbook on Semidefinite, Conic and Polynomial Optimization*, pages 715–754. Springer US, 2012.

- [29] Levent Tunçel and Henry Wolkowicz. Strong duality and minimal representations for cone optimization. *Computational Optimization and Applications*, 53(2):619–648, 2012.
- [30] Hayato Waki. How to generate weakly infeasible semidefinite programs via Lasserre’s relaxations for polynomial optimization. *Optimization Letters*, 6(8):1883–1896, December 2012.
- [31] Hayato Waki and Masakazu Muramatsu. Facial reduction algorithms for conic optimization problems. *Journal of Optimization Theory and Applications*, 158(1):188–215, 2013.
- [32] Hayato Waki, Maho Nakata, and Masakazu Muramatsu. Strange behaviors of interior-point methods for solving semidefinite programming problems in polynomial optimization. *Computational Optimization and Applications*, 53(3):823–844, 2012.

A Reformulation of SDPs

In this section we discuss and prove Proposition 9. So, let $\mathcal{A} : \mathcal{S}^n \rightarrow \mathbb{R}^m$, $b \in \mathbb{R}^m$, $c \in \mathcal{S}^n$ and \mathcal{F} a face of \mathcal{S}_+^n such that $(\mathcal{F}, \mathcal{A}, b, c)$ is both primal and dual strongly feasible. Note that $(\mathcal{F}, \mathcal{A}, b, c)$ is not exactly an SDP, but after appropriate reformulation we can reveal its SDP nature. Recall that due to Proposition 7 there is a $n \times n$ orthogonal matrix q and $r \leq n$ such that

$$q^T \mathcal{F} q = \left\{ \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \in \mathcal{S}^n \mid a \in \mathcal{S}_+^r \right\}. \quad (18)$$

Using the fact that q is orthogonal, we also have

$$q^T \mathcal{F}^* q = \left\{ \begin{pmatrix} a & b \\ b & d \end{pmatrix} \in \mathcal{S}^n \mid a \in \mathcal{S}_+^r \right\}. \quad (19)$$

Now, consider the map $\psi : \mathcal{S}^n \rightarrow \mathcal{S}^r$ that takes $x \in \mathcal{S}^n$ to $\pi_r(q^T x q)$. Note that, restricted to \mathcal{F} , ψ is a bijection and maps \mathcal{F} to \mathcal{S}_+^r . Now, let us compare $\mathcal{F}_P = \{x \in \mathcal{S}^n \mid \mathcal{A}x = b, x \in \mathcal{F}\}$ and $\mathcal{F}_D = \{y \in \mathbb{R}^m \mid c - \mathcal{A}^T y \in \mathcal{F}^*\}$ with the feasible sets of the following pair of primal and dual problems in standard form.

$$\begin{aligned} \inf_x \quad & \langle \psi(c), x \rangle & (P_\psi) \\ \text{subject to} \quad & \langle \psi(\mathcal{A}_i), x_{\mathcal{F}} \rangle = b_i \quad i = 1, \dots, m \\ & x_{\mathcal{F}} \in \mathcal{S}_+^r \end{aligned}$$

$$\begin{aligned} \sup_y \quad & \langle b, y \rangle & (D_\psi) \\ \text{subject to} \quad & \psi(c) - \sum_{i=1}^m \psi(\mathcal{A}_i) y \in \mathcal{S}_+^r, \end{aligned}$$

Note that (P_ψ) and (D_ψ) are *bona fide* SDPs, so as long as they are both strongly feasible, we may invoke (IPO) to solve it. Denote the feasible region of (P_ψ) by $\widehat{\mathcal{F}_P}$ and the feasible region of (D_ψ) by $\widehat{\mathcal{F}_D}$.

Proposition 19. *Consider the restriction of ψ to \mathcal{F} . We have $\psi(\mathcal{F}_P) = \widehat{\mathcal{F}_P}$ and $\psi^{-1}(\widehat{\mathcal{F}_P}) = \mathcal{F}_P$. Moreover, if $x \in \mathcal{F}$, then $\langle c, x \rangle = \langle \psi(c), \psi(x) \rangle$. At the dual side, we have $\widehat{\mathcal{F}_D} = \mathcal{F}_D$.*

Proof. Note that $x \in \mathcal{F}_P$ if and only if $x \in \mathcal{F}$ and $\langle \mathcal{A}_i, x \rangle = b_i$ for every i . However, since q is an orthogonal map, the equality $\langle \mathcal{A}_i, x \rangle = b_i$ holds if and only if $\langle q^T \mathcal{A}_i q, q^T x q \rangle = b_i$ holds. Since $x \in \mathcal{F}$ and (18) holds, we have that $\langle q^T \mathcal{A}_i q, q^T x q \rangle = \langle \psi(\mathcal{A}_i), \psi(x) \rangle$. This same argument also shows that $\langle c, x \rangle = \langle \psi(c), \psi(x) \rangle$. Using the fact $\psi(\mathcal{F}) = \mathcal{S}_+^r$, it follows that \mathcal{F}_P is mapped to $\widehat{\mathcal{F}_P}$ by ψ . On the other hand, if we have some $x_{\mathcal{F}} \in \widehat{\mathcal{F}_P}$, we can consider the unique element $x \in \mathcal{F}$ such that $\psi(x) = x_{\mathcal{F}}$ and we will have $x \in \mathcal{F}_P$.

At the dual side, Equation (19) readily shows that $\widehat{\mathcal{F}_D} = \mathcal{F}_D$. □

Since we are under the assumption that $(\mathcal{F}, \mathcal{A}, b, c)$ is both primal and dual strongly feasible, it follows readily that (P_ψ) and (D_ψ) are strongly feasible as well. So we may invoke (IPO) to solve it. Now, Proposition 19 implies that the optimal value is the same and we may use ψ to recover solutions for (P) and (D). For example, if \hat{x} is an optimal solution for (P_ψ) , then $q\begin{pmatrix} \hat{x} & 0 \\ 0 & 0 \end{pmatrix}q^T$ will be a primal optimal solution for $(\mathcal{F}, \mathcal{A}, b, c)$.